大数定律具体是个什么概念？

要理解大数定律，就必然先要理解小数定律。我会从下面3个方面聊聊：

1）什么是小数定律？

2）什么是大数定律？

3）小数定律和大数定律的动态演示案例

1.什么是小数定律？

喜欢总结规律是人类的天性。例如人们抱着游戏或者认真的态度总结了世界杯足球赛的各种“定律”。“巴西队的礼物”就是其中最著名的定律之一。

“巴西队的礼物”是指：只要巴西夺冠，下一届的冠军就将是主办大赛的东道主，除非巴西队自己夺冠。

我们看下数据：1962年，巴西夺冠，4年后英格兰本土称雄。1970年巴西三夺金杯，1974年轮到东道主西德捧杯。1994年巴西在美国夺冠，1998年法国在本土一鼓作气首次捧得大力神杯。1958年，巴西在瑞典夺取了冠军，4年后他们成功卫冕，收回了“礼物”。

看起来这个定律很有规律，但是这一定律在2006年被破解。

2006年在德国的世界杯，巴西队和东道主德国国家足球队都没能夺冠。

还有还有一些未被破解的定律。这些看似没有道理的神奇定律，之所以神奇，是因为纯属巧合。

世界杯总每四年举办一次，总共才进行了20多届。只要数据足够少，我们总能发现一些没有破解的规律。

如果数据少，随机现象可以看上去很不随机。甚至非常整齐，感觉好像真有规律一样。

iPod最早推出“随机播放”功能的时候，用户发现有些歌曲会被重复播放，他们据此认为播放根本不随机。

苹果公司只好放弃真正的随机算法，用乔布斯本人的话说，就是改进以后的算法使播放“更不随机以至于让人感觉更随机”。

那么，什么是小数定律呢？

小数定律是说，如果统计数据很少，那么事件就表现为各种极端情况，而这些情况都是偶然事件，跟它的期望值一点关系都没有。

统计概率思维让你更值钱

小数定律

如果统计数据很少，那么事件
就表现为各种极端情况，
而这些情况都是偶然事件，
跟它的期望值一点关系都没有。

比如人们知道掷硬币的概率是两面各50% ，于是在连续掷出5个正面之后，却倾向于判断下一次出现反面的几率较大。

再比如 一个只有二十人的乡村中学某年突然有两人考上清华，跟一个有两千人的中学每年都有两百人考上清华，完全没有可比性。

这就好比盲人摸象，你看到的只是偶然事件。

所以，理解“小数定律”最大的一个好处就是你不会再轻易地大惊小怪了，也不会有个人偏见了。

如果统计数据不够大，就什么也说明不了。

正因为如此，我们才不能只凭自己的经验，哪怕加上家人和朋友的经验，去对事物做出判断。我们的经验非常有限。别看个例，看大规模统计。

小数定律里的“跟它的期望值一点关系都没有”，这里的期望值就是后面我会聊到的“大数定律”。

2. 什么是大数定律？

大数定律是我们从统计数字中推测真相的理论基础。

大数定律说如果统计数据足够大，那么事物出现的频率就能无限接近他的期望值。

那么，什么是事物的期望呢？

在我们的生活中，期望是你希望一件事情预期达到什么样的效果。例如，你去面试，期望的薪水是1万5。

在统计概率里，期望也是一样的含义，表示的也是事件未来的预期值，只不过是用更科学的方式来计算出这个数值。

某个事件的期望值，也就是收益，实际上是所有不同结果的和，其中每个结果都是由各自的概率和收益相乘而来。

到这里，你会说：猴子，请说人话。

我还是用例子来说明如何计算期望。

假设你参与了一个掷骰子的游戏，游戏规则是掷出1点可以获得1元，掷出2点可以获得2元，掷出3点可以获得3元，以此类推。

那么在这个游戏中，掷一次骰子的期望值是多少？

前面的概率计算我们已经抛筛子有6种结果，每一个结果的概率都是1/6，因此期望值为：

> 1/6*1元+1/6*2元+1/6*3美元）+1/6*4元+1/6*5元+1/6*6元 =3.5元

统计概率思维让你更值钱

期望=预期值

点数	1	2	3	4	5	6
收益	1元	2元	3元	4元	5元	6元
概率	1/6	1/6	1/6	1/6	1/6	1/6

期望=E(x)=1/6*1元+1/6*2元+1/6*3元+1/6*4元+1/6*5元+1/6*6元=3.5元

微信公众号：猴子聊人物

这个期望3.5元代表什么意思呢？

也就是说只要你持续玩下去，你每次游戏的预期收益是是3.5元。

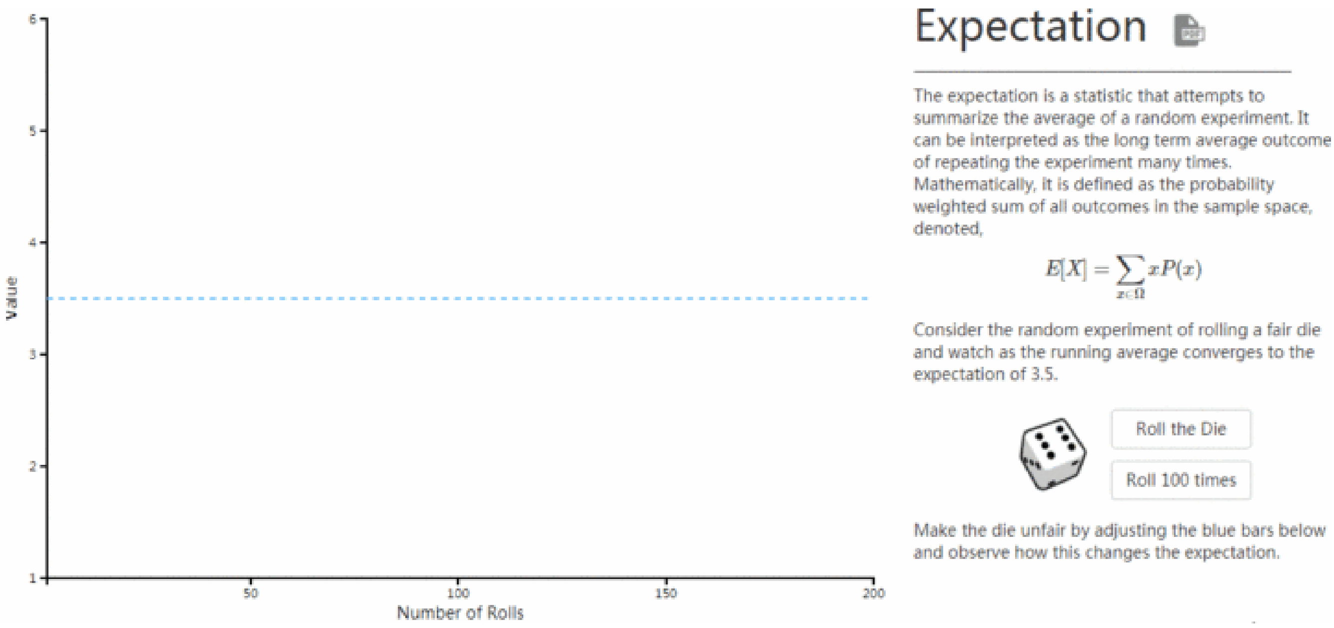
可能你某一次抛筛子赢了1元，某一次抛筛子赢了6元，但是长期来看，你平均下来每次的收益会是3.5元。

其实，通过刚才计算期望的公式，我们可以看出来：**期望的本质是概率的平均值。**

3.小数定律和大数定律的动态演示

下面是前面抛筛子游戏的动态演示过程，横轴表示抛筛子次数，纵轴表示期望值。

你会发现当抛筛子次数少数，期望波动很大。但是当你抛筛子次数大于60次后，就会越来越接近它的期望值3.5。

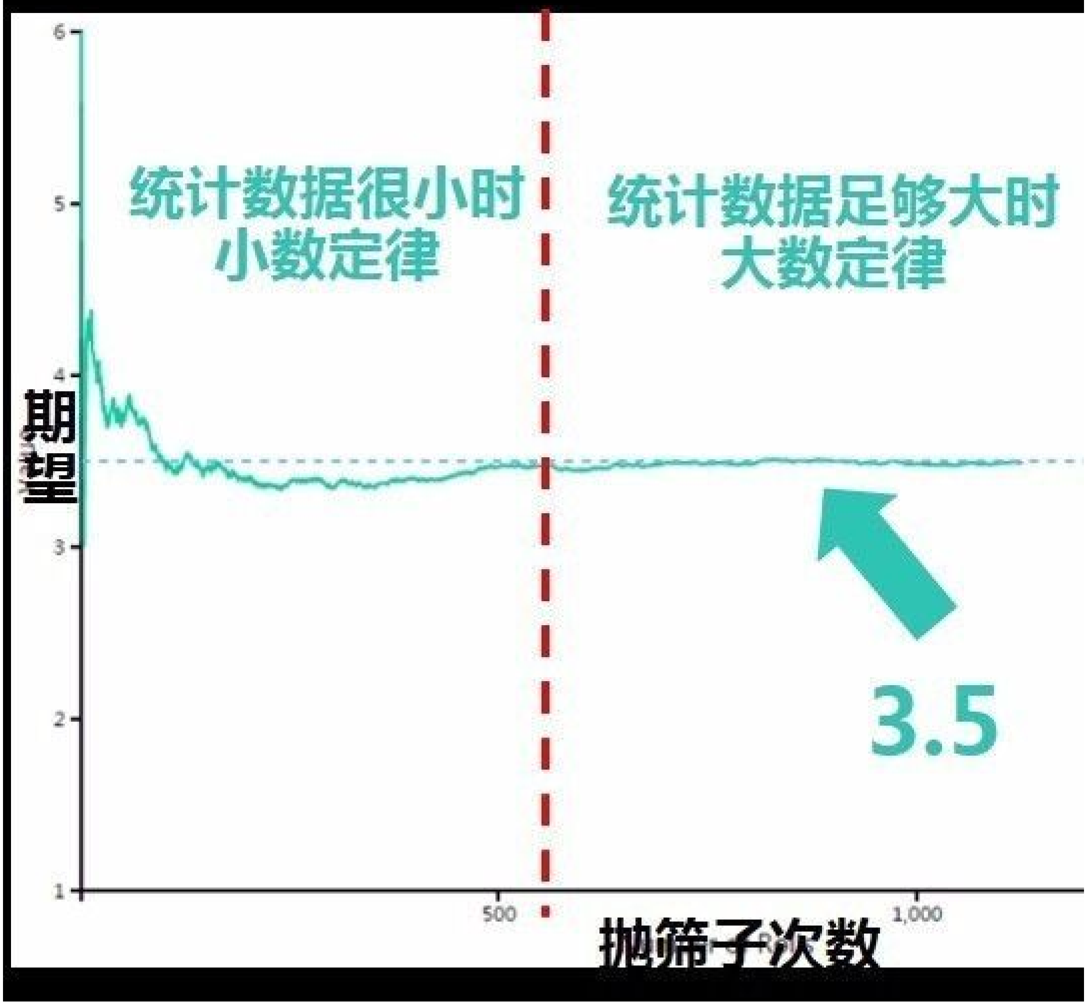


大数定律说如果统计数据足够大，那么事物出现的频率就能无限接近他的期望。 这里的图片是刚才扔100次筛子的期望图，横轴是扔筛子的次数，纵轴是筛子抛出点数的期望。

统计概率思维让你更值钱

大数定律

如果统计数据足够大，那么事物出现的频率就能无限接近他的期望



我们看到当在扔筛子次数很少，少于60次的情况下，抛出点数波动很大。这就是小数定律，如果统计数据很少，那么事件就表现为各种极端情况，而这些情况都是偶然事件，跟它的期望值一点关系都没有。 但是随着抛出次数的增加，最终抛出点数的期望是3.5。和我们刚才计算出的期望数值的一样的。

粗略看一下，抛出筛子点数是3.5的期望似乎是一个无效数据，毕竟筛子每一面的数值是整数，而不是小数。

但是，期望是一个非常有用的参考数据，就像刚才我们说的抛筛子游戏，你可以通过比较成本投入和期望收益，你就能知道做这件事是不是“值得”。

如果在上述游戏中，每掷一次骰子需要缴纳5元，你还玩吗？

根据大数据定律，如果统计数据足够大，那么事物出现的频率就能无限接近他的期望，也就是只要你一直持续玩下去，你每次游戏的预期收益是3.5元。

如果你玩1万次这个游戏，投入成本5元乘以1万次是5万元，而预期收益3.5元乘以1万次是3.5万元，那么你将输掉的是5万元-3.5万元等于1.5万元。这种赔本的游戏你还敢玩吗？

这也可以解释为什么买彩票，长期来看是一种赔本的游戏。彩票从本质上来看其实就是一种赌博行为。我们也可以用大数定律来解释为什么赌场从长期来看总是挣钱的问题。

赌场内所有项目的概率都是有利于赌场老板的（出“老千”的赌客不考虑在内）。如果赌场的营业时间足够长，吸引的下注人数也足够多，那么赌场从赌桌赚到的钱肯定要比付出的要多。